J.1.4.4.3 The Archive Benchmark for Workstreams 4, 5 and 6

The archive benchmark will measure the time to retrieve a sample dataset comprised of files representing the Government's file size distribution from the HSMS. The sample dataset will be comprised of 70 "large" files approximately 4dGB in size and 120 "small" files approximately 500dMB in size.

The sample dataset will be retrieved from the HSMS to disk storage accessible by the analysis applications for workstreams 4, 5 and 6. At the initial delivery date for workstreams 4, 5 and 6, retrieval of the sample dataset shall take less than 20 minutes.

J.1.4.4.4 The Archive Benchmark for Workstreams 7, 8 and 9

The archive benchmark will measure the time to retrieve a sample dataset comprised of files representing the Government's file size distribution from the HSMS. The sample dataset will be comprised of 60 "large" files approximately 4dGB in size and 200 "small" files approximately 400dMB in size.

The sample dataset will be retrieved from the HSMS to disk storage accessible by the analysis applications for workstreams 7, 8 and 9. At the initial delivery date for workstreams 7, 8 and 9, retrieval of the sample dataset shall take less than 150 minutes. At midlife upgrade for workstreams 7, 8, and 9, retrieval of the sample data set shall take less than 100 minutes.

J.1.4.5 Benchmark Model Overview

J.1.4.5.1 Workstream 1: CM2-ESM

The CM2 Earth System Model (ESM) is comprised of the N45L24 bgrid atmosphere core (i.e., 144 x 90 horizontal resolution with 24 levels) with land and ice model components coupled to a one-degree MOM4 ocean model. While the atmosphere portion of the model is malleable with respect to layout and PE count, the best performance of the current production model is achieved with a STATIC_MEMORY MOM4. Thus, a given executable may run multiple atmosphere configurations, but only one ocean layout. For example, the same executable may be used to run on 120 and 150 PEs in 60atm+60ocn or 90atm+60ocn concurrent mode; serial mode will always require a unique executable assuming that STATIC_MEMORY MOM4 shows performance advantages over the malleable form.

Multiple sample PE configurations have been provided. Concurrent mode examples carry the designation of the ocean portion of the model. Thus, cm2.30, cm2.60, cm2.90, cm2.120, cm2.150 and cm2.180 are all serial mode examples. Test cases cm2.30o.c (60PEs), cm2.60o.c (120 and 150PEs), cm2.72o.c (180PEs), cm2.80o.c (200PEs) and cm2.90o.c (180PEs) are all concurrent cases. All but cm2.72o.c and cm2.80o.c use executables that also run for the serial cases; the atmosphere run on 72 or 80PEs is not interesting on current architectures and so these executables have not been run in serial mode.

One of the goals of a concurrent mode configuration shall be to load balance between the ocean model and remaining components. Until recently, the 60+60 and 90+90 configurations provided fairly good balance. Improvements in the time-stepping scheme for MOM4 have just been introduced that change this balance. Moreover, it's expected the port to different architectures will produce different performance features for each of the model components. Thus, finding the best balance of processing element (PE) configurations will be part